BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

As rescanning documents *will not* correct images, please do not report the images to the Image Problem Mailbox.

PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 :
C12P 21/06, C12N 5/00, 15/00, C07H 21/02

A1

(11) International Publication Number: WO 99/28492

(43) International Publication Date: 10 June 1999 (10.06.99)

(21) International Application Number: PCT/US98/25497

(22) International Filing Date: 2 December 1998 (02.12.98)

(30) Priority Data:

08/984,288	3 December 1997 (03.12.97)	US
60/073,747	5 February 1998 (05.02.98)	US
90/060,504	15 April 1998 (15.04.98)	US

(81) Designated States: JP, European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).

Published
With international search report.

(71) Applicant: SMITHKLINE BEECHAM CORPORATION [US/US]; One Franklin Plaza, Philadelphia, PA 19103 (US).

(72) Inventors: AMES, Robert, S., Jr.; 276 Ellis Road, Havertown, PA 19083 (US). SARAU, Henry, M.; 348 Maple Avenue, Harleysville, PA 19438 (US). FOLEY, James, J.; 117 Timothy Circle, Radnor, PA 19087 (US). BERGSMA, Derk; 271 Irish Road, Berwyn, PA 19312 (US). ELLIS, Catherine, E.; 831 Fordham Place, Glassboro, NJ 08028 (US). CHAMBERS, Jon, K.; 32 Badcock Road, Haslingfield, Cambridge (GB).

(74) Agents: ANDERSEN, Robert, L. et al.; Ratner & Prestia, Suite 301, One Westlakes (Berwyn), P.O. Box 980, Valley Forge, PA 19482-0980 (US).

(54) Title: A METHOD OF FINDING AGONIST AND ANTAGONIST TO HUMAN 11CB SPLICE VARIANT

(57) Abstract

Human 11cb splice variant polypeptides and DNA (RNA) encoding such an 11cb splice variant and a procedure for producing such polypeptides by recombinant techniques are disclosed. Also disclosed are methods for utilizing such an 11cb splice variant for the treatment of to treat infections, such as bacterial, fungal, protozoan and viral infections, particularly infection caused by HIV-1 or HIV-2; pain; cancers; diabetes; obesity; feeding and drinking abnormalities, such as anorexia and bulimia; asthma; Parkinson's disease; both acute and congestive heart failure; hypotension; hypertension; urinary retention; osteoporosis; angina pectoris; myocardial infarction; ulcers; allergies; benign prostatic hypertrophy and psychotic and neurological disorders, including anxiety, schizophrenia, manic depression, delirium, dementia or severe mental retardation, and dyskinesias, such as Huntington's disease or Gilles de la Tourett's syndrome; among others. Antagonists against such an 11cb splice variant and their use as a therapeutic to treat infections, such as bacterial, fungal, protozoan and viral infections, particularly infection caused by HIV-1 or HIV-2; pain; cancers; diabetes; obesity; feeding and drinking abnormalities, such as anorexia and bulimia; asthma; Parkinson's disease; both acute and congestive heart failure; hypotension; hypertension; urinary retention; osteoporosis; angina pectoris; myocardial infarction; ulcers; allergies; benign prostatic hypertrophy and psychotic and neurological disorders, including anxiety, schizophrenia, manic depression, delirium, dementia or severe mental retardation, and dyskinesias, such as Huntington's disease or Gilles de la Tourett's syndrome; among others, are also disclosed. Also disclosed are diagnostic assays for detecting diseases related to mutations in the nucleic acid sequences and altered concentrations of the polypeptides. Also disclosed are diagnostic assays for detecting mutations in the polynucleotides encoding the 11cb splice variant and for detecting altered levels of the polypeptide in a host.

A METHOD OF FINDING AGONIST AND ANTAGONIST TO HUMAN 11CB SPLICE VARIANT

FIELD OF THE INVENTION

This invention relates, in part, to newly identified polynucleotides and polypeptides; variants and derivatives of the polynucleotides and polypeptides; processes for making the polynucleotides and the polypeptides, and their variants and derivatives; agonists and antagonists of the polypeptides; and uses of the polynucleotides, polypeptides, variants, derivatives, agonists and antagonists. In particular, in these and in other regards, the invention relates to polynucleotides and polypeptides of the human 11cb splice variant.

BACKGROUND OF THE INVENTION

This invention relates to newly identified polynucleotides, polypeptides encoded by such polynucleotides, the use of such polynucleotides and polypeptides, as well as the production of such polynucleotides and polypeptides. More particularly, the polypeptides of the present invention are human 7-transmembrane receptors. The invention also relates to inhibiting or activating the action of such polypeptides.

It is well established that many medically significant biological processes are mediated by proteins participating in signal transduction pathways that involve G-proteins and/or second messengers, *e.g.*, cAMP (Lefkowitz, *Nature*, (1991) 351: 353-354). Herein these proteins are referred to as proteins participating in pathways with G-proteins or PPG proteins. Some examples of these proteins include the GPC receptors, such as those for adrenergic agents and dopamine (Kobilka, B.K., *et al.*, *PNAS*, (1987), 84: 46-50; *Kobilka, B.K., et al., Science*, (1987), 238: 650-656; Bunzow, J.R., *et al., Nature*, (1988), 336:783-787), G-proteins themselves, effector proteins, *e.g.*, phospholipase C, adenyl cyclase, and phosphodiesterase, and actuator proteins, *e.g.*, protein kinase A and protein kinase C (Simon, M.I., *et al., Science*, 252: 802-8 (1991)).

For example, in one form of signal transduction, the effect of hormone binding is activation of the enzyme, adenylate cyclase, inside the cell. Enzyme activation by hormones is dependent on the presence of the nucleotide GTP. GTP also influences hormone binding. A G-protein connects the hormone receptor to adenylate cyclase. G-protein was shown to exchange GTP for bound GDP when activated by a hormone receptor. The GTP-carrying form then binds to activated adenylate cyclase. Hydrolysis of GTP to GDP, catalyzed by the G-protein itself, returns the G-protein to its basal, inactive form. Thus, the G-protein serves a dual role, as an

serines, a TM6 asparagine and TM6 or TM7 phenylalanines or tyrosines are also implicated in ligand binding.

G-protein coupled receptors can be intracellularly coupled by heterotrimeric G-proteins to various intracellular enzymes, ion channels and transporters. *See* Johnson, *et al., Endoc. Rev.*, (1989) 10: 317-331). Different G-protein a-subunits preferentially stimulate particular effectors to modulate various biological functions in a cell. Phosphorylation of cytoplasmic residues of G-protein coupled receptors have been identified as an important mechanism for the regulation of G-protein coupling of some G-protein coupled receptors. G-protein coupled receptors are found in numerous sites within a mammalian host.

Since, over the past 15 years, nearly 150 therapeutic agents targeting 7 transmembrane (7 TM) receptors have been successfully introduced onto the market. This indicates that these receptors have an established, proven history as therapeutic targets. Clearly, there is a need for identification and characterization of further receptors which can play a role in preventing, ameliorating or correcting dysfunctions or diseases, including, but not limited to, infections such as bacterial, fungal, protozoan and viral infections, particularly infection caused by HIV-1 or HIV-2; pain; cancers; diabetes; obesity; feeding and drinking abnormalities, such as anorexia and bulimia; asthma; Parkinson's disease; both acute and congestive heart failure; hypotension; hypertension; urinary retention; osteoporosis; angina pectoris; myocardial infarction; ulcers; allergies; benign prostatic hypertrophy and psychotic and neurological disorders, including anxiety, schizophrenia, manic depression, delirium, dementia or severe mental retardation, and dyskinesias, such as Huntington's disease or Gilles dela Tourett's syndrome, among others.

The polypeptide of the present invention has the conserved 7 transmembrane residues, and have amino acid sequence homology to known G-protein couples receptors.

The orginal human 11cb clone has been previously disclosed by Applicants in PCT WO 96/18651, published June 20, 1996.

SUMMARY OF THE INVENTION

Toward these ends, and others, it is an object of the present invention to provide polypeptides, *inter alia*, that have been identified as a novel human 11cb splice variant by homology between the amino acid sequence set out in Figure 1 (SEQ ID NO: 2) and known amino acid sequences of other proteins such as mouse cDNA, rat calcitonin receptor A, rat calcitonin receptor B, and hormone receptor EMR1.

In accordance with another aspect of the present invention there are provided methods of screening for compounds which bind to and activate (agonist) or inhibit activation (antagonist) of the receptor polypeptides of the present invention and for receptor ligands.

In particular, the preferred method for identifying agonist or antagonist of a receptor of the present invention comprises:

contacting a cell expressing on the surface thereof the receptor, said receptor being associated with a second component capable of providing a detectable signal in response to the binding of a compound to said receptor, with a compound to be screened under conditions to permit binding to the receptor; and

determining whether the compound binds to and activates or inhibits the receptor by measuring the level of a signal generated from the interaction of the compound with the receptor.

In a further preferred embodiment, the method further comprises conducting the identification of agonist or antagonist in the presence of labeled or unlabeled MCH.

In another embodiment of the method for identifying agonist or antagonist of a receptor of the present invention comprises:

determining the inhibition of binding of a ligand to cells which have the receptor on the surface thereof, or to cell membranes containing the recetpor, in the presence of a candidate compound under conditions to permit binding to the receptor, and determining the amount of ligand bound to the receptor, such that a compound capable of causing reduction of binding of a ligand is an agonist or antagonist. Preferably the ligand is MCH. Yet more preferably MCH is labeled.

It is another object of the invention to provide a process for producing the aforementioned polypeptides, polypeptide fragments, variants and derivatives, fragments of the variants and derivatives, and analogs of the foregoing. In a preferred embodiment of this aspect of the invention there are provided methods for producing the aforementioned 11cb splice variant polypeptides comprising culturing host cells having expressibly incorporated therein an exogenously-derived 11cb splice variant-encoding polynucleotide under conditions for expression of 11cb splice variant in the host and then recovering the expressed polypeptide.

In accordance with another object the invention there are provided products, compositions, processes and methods that utilize the aforementioned polypeptides and polynucleotides for research, biological, clinical and therapeutic purposes, *inter alia*.

modulating ligand binding, due to their expected biological properties, which may be used in diagnostic, therapeutic and/or research applications.

It is still another object of the present invention to provide synthetic, isolated or recombinant polypeptides which are designed to inhibit or mimic various 11cb splice variants or fragments thereof, as receptor types and subtypes.

In accordance with certain preferred embodiments of this and other aspects of the invention there are provided probes that hybridize to human 11cb splice variant sequences.

In certain additional preferred embodiments of this aspect of the invention there are provided antibodies against 11cb splice variant polypeptides. In certain particularly preferred embodiments in this regard, the antibodies are highly selective for human 11cb splice variant.

In accordance with another aspect of the present invention, there are provided 11cb splice variant agonists. Among preferred agonists are molecules that mimic the 11cb splice variant, that bind to 11cb splice variant-binding molecules or receptor molecules, and that elicit or augment 11cb splice variant-induced responses. Also among preferred agonists are molecules that interact with 11cb splice variant or 11cb splice variant polypeptides, or with other modulators of 11cb splice variant activities, and thereby potentiate or augment an effect of 11cb splice variant or more than one effect of 11cb splice variant.

In accordance with yet another aspect of the present invention, there are provided 11cb splice variant antagonists. Among preferred antagonists are those which mimic the 11cb splice variant so as to bind to the 11cb splice variant receptor or binding molecules but not elicit an 11cb splice variant-induced response or more than one 11cb splice variant-induced response. Also among preferred antagonists are molecules that bind to or interact with the 11cb splice variant so as to inhibit an effect of 11cb splice variant or more than one effect of 11cb splice variant or which prevent expression of 11cb splice variant.

In a further aspect of the invention there are provided compositions comprising an 11cb splice variant polynucleotide or an 11cb splice variant polypeptide for administration to cells *in vitro*, to cells *ex vivo* and to cells *in vivo*, or to a multicellular organism. In certain particularly preferred embodiments of this aspect of the invention, the compositions comprise an 11cb splice variant polynucleotide for expression of an 11cb splice variant polypeptide in a host organism for treatment of disease. Particularly preferred in this regard is expression in a human patient for treatment of a dysfunction associated with aberrant endogenous activity of the 11cb splice variant.

natural state is not "isolated," but the same polynucleotide or polypeptide separated from the coexisting materials of its natural state is "isolated", as the term is employed herein.

"Polynucleotide(s)" generally refers to any polyribonucleotide or polydeoxribonucleotide, which may be unmodified RNA or DNA or modified RNA or DNA. Polynucleotides as used herein refers to, among others, single- and double-stranded DNA, DNA that is a mixture of single- and double-stranded regions, single- and double-stranded RNA, and RNA that is mixture of single- and double-stranded regions, hybrid molecules comprising DNA and RNA that may be single-stranded or, more typically, double-stranded or a mixture of single- and double-stranded regions. In addition, polynucleotide as used herein refers to triple-stranded regions comprising RNA or DNA or both RNA and DNA. The strands in such regions may be from the same molecule or from different molecules. The regions may include all of one or more of the molecules, but more typically involve only a region of some of the molecules. The term polynucleotide also includes DNAs or RNAs containing one or more modified bases. Thus, DNAs or RNAs with backbones modified for stability or for other reasons are polynucleotides as that term is intended herein. Moreover, DNAs or RNAs comprising unusual bases, such as inosine, or modified bases, such as tritylated bases, to name just two examples, are polynucleotides. It will be appreciated that a great variety of modifications have been made to DNA and RNA that serve many useful purposes known to those of skill in the art. The term polynucleotide, as it is employed herein, embraces such chemically, enzymatically or metabolically modified forms of polynucleotides, as well as the chemical forms of DNA and RNA characteristic of viruses and cells, including *inter alia* simple and complex cells. The term polynucleotide, as used herein, also embraces relatively short polynucleotides, often referred to as oligonucleotides.

"Polypeptides", as used herein, includes all polypeptides as described below. The basic structure of polypeptides is well known and has been described in the art. The term is used herein to refer to any peptide or protein comprising two or more amino acids joined to each other by peptide bonds. As used herein, the term refers to both short chains, which also commonly are referred to in the art as peptides, oligopeptides and oligomers, for example, and to longer chains, which generally are referred to in the art as proteins, of which there are many types. Polypeptides often contain amino acids other than the 20 amino acids commonly referred to as the 20 naturally occurring amino acids, and that many amino acids, including the terminal amino acids, may be modified in a given polypeptide, either by natural processes, such as processing and other post-translational modifications, or by chemical modification techniques which are well

present invention, as well. For instance, the amino terminal residue of polypeptides made in *E. coli*, prior to processing, almost invariably will be N-formylmethionine.

The modifications that occur in a polypeptide often will be a function of how it is made. For polypeptides made by expressing a cloned gene in a host, for instance, the nature and extent of the modifications in large part will be determined by the host cell's posttranslational modification capacity and the modification signals present in the polypeptide amino acid sequence. For instance, as is well known, glycosylation often does not occur in bacterial hosts such as *E. coli*. Accordingly, when glycosylation is desired, a polypeptide should be expressed in a glycosylating host, generally a eukaryotic cell. Insect cells often carry out the same posttranslational glycosylations as mammalian cells and, for this reason, insect cell expression systems have been developed to express efficiently mammalian proteins having the native patterns of glycosylation, *inter alia*. Similar considerations apply to other modifications.

It will be appreciated that the same type of modification may be present in the same or varying degrees at several sites in a given polypeptide. Also, a given polypeptide may contain many types of modifications.

The term polypeptide encompasses all such modifications, particularly those that are present in polypeptides synthesized by expressing a polynucleotide in a host cell.

"Variant(s)," as the term is used herein, are polynucleotides or polypeptides that differ from a reference polynucleotide or polypeptide respectively. Variants in this sense are described below and elsewhere in the present disclosure in greater detail. (1) A polynucleotide that differs in nucleotide sequence from another, reference polynucleotide. Changes in the nucleotide sequence of the variant may be silent, *i.e.*, they may not alter the amino acids encoded by the polynucleotide. Where alterations are limited to silent changes of this type a variant will encode a polypeptide with the same amino acid sequence as the reference polypeptide. Changes in the nucleotide sequence of the variant may alter the amino acid sequence of a polypeptide encoded by the reference polynucleotide. Such nucleotide changes may result in amino acid substitutions, additions, deletions, fusions and truncations in the polypeptide encoded by the reference sequence, as discussed below. (2) A polypeptide that differs in amino acid sequence from another, reference polypeptide. Generally, differences are limited so that the sequences of the reference and the variant are closely similar overall and, in many region, identical. A variant and reference polypeptide may differ in amino acid sequence by one or more substitutions, additions, deletions, fusions and truncations, which may be present in any combination. (3) A variant may

polynucleotide sequences, as the case may be, as determined by the match between strings of such sequences. "Similarity" between two polypeptides is determined by comparing the amino acid sequence and its conserved amino acid substitutes of one polypeptide to the sequence of a second polypeptide. "Identity" and "similarity" can be readily calculated by known methods, including but not limited to those described in (*Computational Molecular Biology*, Lesk, A.M., ed., Oxford University Press, New York, 1988; *Biocomputing: Informatics and Genome Projects*, Smith, D.W., ed., Academic Press, New York, 1993; *Computer Analysis of Sequence Data*, Part I, Griffin, A.M., and Griffin, H.G., eds., Humana Press, New Jersey, 1994; *Sequence Analysis in Molecular Biology*, von Heinje, G., Academic Press, 1987; and *Sequence Analysis Primer*, Gribskov, M. and Devereux, J., eds., M Stockton Press, New York, 1991; and Carillo, H., and Lipman, D., SIAM *J. Applied Math., 48:* 1073 (1988). Preferred methods to determine identity are designed to give the largest match between the sequences tested. Methods to determine identity and similarity are codified in publicly available computer programs. Preferred computer program methods to determine identity and similarity between two sequences include, but are not limited to, the GCG program package (Devereux, J., et al., *Nucleic Acids Research 12(1):* 387 (1984)), BLASTP, BLASTN, and FASTA (Atschul, S.F. et al., *J. Molec. Biol. 215:* 403-410 (1990). The BLAST X program is publicly available from NCBI and other sources (*BLAST Manual*, Altschul, S., *et al.*, NCBI NLM NIH Bethesda, MD 20894; Altschul, S., *et al., J. Mol. Biol. 215*: 403-410 (1990). The well known Smith Waterman algorithm may also be used to determine identity.

Preferred parameters for polypeptide sequence comparison include the following:

1) Algorithm: Needleman and Wunsch, J. Mol Biol. 48: 443-453 (1970)

Comparison matrix: BLOSSUM62 from Hentikoff and Hentikoff, Proc. Natl. Acad. Sci. USA. 89:10915-10919 (1992)

Gap Penalty: 12

Gap Length Penalty: 4

A program useful with these parameters is publicly available as the "gap" program from Genetics Computer Group, Madison WI. The aforementioned parameters are the default parameters for peptide comparisons (along with no penalty for end gaps).

Preferred parameters for polynucleotide comparison include the following:

1) Algorithm: Needleman and Wunsch, J. Mol Biol. 48: 443-453 (1970)

interspersed either individually among the nucleic acids in the reference sequence or in one or more contiguous groups within the reference sequence. The number of nucleic acid alterations for a given percent identity is determined by multiplying the total number of amino acids in SEQ ID NO:2 by the integer defining the percent identity divided by 100 and then subtracting that product from said total number of amino acids in SEQ ID NO:2, or: $n_n = x_n - (x_n \cdot y)$, wherein n_n is the number of amino acid alterations, x_n is the total number of amino acids in SEQ ID NO:2, y is, for instance 0.70 for 70%, 0.80 for 80%, 0.85 for 85% etc., , is the symbol for the multiplication operator, and wherein any non-integer product of x_n and y is rounded down to the nearest integer prior to subtracting it from x_n.

Preferred polypeptide embodiments further include an isolated polypeptide comprising a polypeptide having at least a 50,60, 70, 80, 85, 90, 95, 97 or 100% identity to a polypeptide reference sequence of SEQ ID NO:2, wherein said polypeptide sequence may be identical to the reference sequence of SEQ ID NO: 2 or may include up to a certain integer number of amino acid alterations as compared to the reference sequence, wherein said alterations are selected from the group consisting of at least one amino acid deletion, substitution, including conservative and non-conservative substitution, or insertion, and wherein said alterations may occur at the amino- or carboxy-terminal positions of the reference polypeptide sequence or anywhere between those terminal positions, interspersed either individually among the amino acids in the reference sequence or in one or more contiguous groups within the reference sequence, and wherein said number of amino acid alterations is determined by multiplying the total number of amino acids in SEQ ID NO:2 by the integer defining the percent identity divided by 100 and then subtracting that product from said total number of amino acids in SEQ ID NO:2, or: $n_a \quad x_a - (x_a \quad y)$, wherein n_a is the number of amino acid alterations, x_a is the total number of amino acids in SEQ ID NO:2, y is 0.50 for 50%, 0.60 for 60%, 0.70 for 70%, 0.80 for 80%, 0.85 for 85%, 0.90 for 90%, 0.95 for 95%, 0.97 for 97% or 1.00 for 100%, and is the symbol for the multiplication operator, and wherein any non-integer product of x_a and y is rounded down to the nearest integer prior to subtracting it from x_a.

By way of example, a polypeptide sequence of the present invention may be identical to the reference sequence of SEQ ID NO:2, that is it may be 100% identical, or it may include up to a certain integer number of amino acid alterations as compared to the reference sequence such that the percent identity is less than 100% identity. Such alterations are selected from the group consisting of at least one amino acid deletion, substitution, including conservative and non-conservative substitution, or insertion, and wherein said alterations may occur at the amino- or

The coding sequence which encodes the polypeptide may be identical over its entire length to the coding sequence of the polynucleotide shown in Figure 1 (SEQ ID NO: 1). It also may be a polynucleotide with a different sequence, which, as a result of the redundancy (degeneracy) of the genetic code, also encodes the polypeptide of Figure 1 (SEQ ID NO: 2).

Polynucleotides of the present invention which encode the polypeptide of Figure 1 (SEQ ID NO: 2) may include, but are not limited to, the coding sequence for the mature polypeptide, by itself; the coding sequence for the mature polypeptide and additional coding sequences, such as those encoding a leader or secretory sequence, such as a pre-, or pro- or prepro- protein sequence; and the coding sequence of the mature polypeptide, with or without the aforementioned additional coding sequences, together with additional, non-coding sequences, including, but not limited to, introns and non-coding 5' and 3' sequences, such as the transcribed, non-translated sequences that play a role in transcription, and mRNA processing, including splicing and polyadenylation signals, for example, for ribosome binding and stability of mRNA. Coding sequences which provide additional functionalities may also be incorporated into the polypeptide. Thus, for instance, the polypeptide may be fused to a marker sequence, such as a peptide, which facilitates purification of the fused polypeptide. In certain preferred embodiments of this aspect of the invention, the marker sequence is a hexa-histidine peptide, such as the tag provided in the pQE vector (Qiagen, Inc.). As described in *Gentz, et al., Proc. Natl. Acad. Sci., USA*, 1989, 86: 821-824, for instance, hexa-histidine provides for convenient purification of the fusion protein. In other embodiment the marker sequence is a HA tag. Many other such tags are commercially abatable.

In accordance with foregoing, the term "polynucleotide encoding a polypeptide" also encompasses polynucleotides that include a single continuous region or discontinuous regions encoding the polypeptide (for example, interrupted by introns) together with additional regions, that also may contain coding and/or non-coding sequences.

The present invention further relates to variants of the polynucleotides which encode for variants of the polypeptide having the deduced amino acid sequence of Figure 1 (SEQ ID NO: 2).

Among particularly preferred embodiments of the invention are polynucleotides encoding polypeptides having the amino acid sequence of the 11cb splice variant set out in Figure 1 (SEQ ID NO: 2) and variants thereof.

Further preferred embodiments are polynucleotides encoding variants of the 11cb splice variant that have the amino acid sequence of the 11cb splice variant polypeptide of Figure 1 (SEQ

Polypeptides

The present invention further relates to a human 11cb splice variant polypeptide which has the deduced amino acid sequence of Figure 1 (SEQ ID NO: 2).

The polypeptide of the present invention may be a recombinant polypeptide, a natural polypeptide or a synthetic polypeptide. In certain preferred embodiments, it is a recombinant polypeptide.

Among the particularly preferred embodiments of the invention are polypeptides having the amino acid sequence of 11cb splice variant, set out in Figure 1 (SEQ ID NO: 2), and variants thereof. Other preferred embodiments of the invention are polypeptides having the amino acid sequence of 11cb splice variant, and variants thereof

Among preferred variants are those that vary from a reference by conservative amino acid substitutions. Such substitutions are those that substitute a given amino acid in a polypeptide by another amino acid of like characteristics. Typically seen as conservative substitutions are the replacements, one for another, among the aliphatic amino acids Ala, Val, Leu and Ile; interchange of the hydroxyl residues Ser and Thr, exchange of the acidic residues Asp and Glu, substitution between the amide residues Asn and Gln, exchange of the basic residues Lys and Arg and replacements among the aromatic residues Phe and Tyr.

Further preferred are variants of the fragments, having the amino acid sequence of the 11cb splice variant polypeptide of Figure 1 (SEQ ID NO: 2), in which several, 5 to 10, 1 to 5, 1 to 3, 2, 1 or no amino acid residues are substituted, deleted or added, in any combination.

Especially preferred among these are silent substitutions, additions and deletions, which do not alter the properties and activities of the 11cb splice variant. Also especially preferred in this regard are conservative substitutions.

Most highly preferred are polypeptides having the amino acid sequence of Figure 1 (SEQ ID NO: 2) without substitutions.

The polypeptides and polynucleotides of the present invention are preferably provided in an isolated form, and preferably are purified to homogeneity.

The polypeptides of the present invention include the polypeptide of SEQ ID NO: 2 (in particular the mature polypeptide) as well as polypeptides which have at least 91% identity to the polypeptide of SEQ ID NO: 2

regard include fragments that comprise alpha-helix and alpha-helix forming regions, beta-sheet and beta-sheet-forming regions, turn and turn-forming regions, coil and coil-forming regions, hydrophilic regions, hydrophobic regions, alpha amphipathic regions, beta amphipathic regions, flexible regions, surface-forming regions, substrate binding region, and high antigenic index regions of the 11cb splice variant, and combinations of such fragments.

Preferred regions are those that mediate activities of the 11cb splice variant. Most highly preferred in this regard are fragments that have a chemical, biological or other activity of the 11cb splice variant, including those with a similar activity or an improved activity, or with a decreased undesirable activity. Further preferred polypeptide fragments are those that are antigenic or immunogenic in an animal, especially in a human.

It will be appreciated that the invention also relates to, among others, polynucleotides encoding the aforementioned fragments, polynucleotides that hybridize to polynucleotides encoding the fragments, particularly those that hybridize under stringent conditions, and polynucleotides, such as PCR primers, for amplifying polynucleotides that encode the fragments. In these regards, preferred polynucleotides are those that correspond to the preferred fragments, as discussed above.

Vectors, host cells, expression

The present invention also relates to vectors which comprise a polynucleotide or polynucleotides of the present invention, and host cells which are genetically engineered with vectors of the invention and the production of polypeptides of the invention by recombinant techniques.

Host cells can be genetically engineered to incorporate polynucleotides and express polypeptides of the present invention. Introduction of a polynucleotides into the host cell can be affected by calcium phosphate transfection, DEAE-dextran mediated transfection, transvection, microinjection, cationic lipid-mediated transfection, electroporation, transduction, scrape loading, ballistic introduction, infection or other methods. Such methods are described in many standard laboratory manuals, such as Davis, *et al.*, BASIC METHODS IN MOLECULAR BIOLOGY, (1986) and Sambrook, *et al.*, MOLECULAR CLONING: A LABORATORY MANUAL, 2nd Ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y. (1989).

Representative examples of appropriate hosts include bacterial cells, such as *streptococci, staphylococci, E. coli, streptomyces* and *Bacillus subtilis* cells; fungal cells, such as yeast cells

prokaryotic and eukaryotic hosts, are well known and employed routinely by those of skill in the art.

A great variety of expression vectors can be used to express a polypeptide of the invention. Such vectors include, among others, chromosomal, episomal and virus-derived vectors, *e.g.*, vectors derived from bacterial plasmids, from bacteriophage, from transposons, from yeast episomes, from insertion elements, from yeast chromosomal elements, from viruses such as baculoviruses, papova viruses, such as SV40, vaccinia viruses, adenoviruses, fowl pox viruses, pseudorabies viruses and retroviruses, and vectors derived from combinations thereof, such as those derived from plasmid and bacteriophage genetic elements, such as cosmids and phagemids, all may be used for expression in accordance with this aspect of the present invention. Generally, any vector suitable to maintain, propagate or express polynucleotides to express a polypeptide in a host may be used for expression in this regard.

The appropriate DNA sequence may be inserted into the vector by any of a variety of well-known and routine techniques, such as, for example, those set forth in Sambrook, *et al.*, MOLECULAR CLONING, A LABORATORY MANUAL, 2nd Ed.; Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York (1989).

The DNA sequence in the expression vector is operatively linked to appropriate expression control sequence(s), including, for instance, a promoter to direct mRNA transcription. Representatives of such promoters include, but are not limited to, the phage lambda PL promoter, the *E. coli* lac, trp and tac promoters, the SV40 early and late promoters and promoters of retroviral LTRs.

In general, expression constructs will contain sites for transcription initiation and termination, and, in the transcribed region, a ribosome binding site for translation. The coding portion of the mature transcripts expressed by the constructs will include a translation initiating codon, for example, AUG or GUG, at the beginning and a termination codon appropriately positioned at the end of the polypeptide to be translated.

In addition, the constructs may contain control regions that regulate as well as engender expression. Generally, in accordance with many commonly practiced procedures, such regions will operate by controlling transcription, such as transcription factors, repressor binding sites and termination, among others.

Vectors for propagation and expression generally will include selectable markers and amplification regions, such as, for example, those set forth in Sambrook et al.

structural sequence, and a selectable marker to permit isolation of vector containing cells after exposure to the vector.

Polynucleotides of the invention, encoding the heterologous structural sequence of a polypeptide of the invention generally will be inserted into the vector using standard techniques so that it is operably linked to the promoter for expression. The polynucleotide will be positioned so that the transcription start site is located appropriately 5' to a ribosome binding site. The ribosome binding site will be 5' to the codon that initiates translation of the polypeptide to be expressed, for example AUG or GUG. Generally, there will be no other open reading frames that begin with an initiation codon, usually AUG, and lie between the ribosome binding site and the initiation codon. Also, generally, there will be a translation stop codon at the end of the polypeptide and there will be a polyadenylation signal in constructs for use in eukaryotic hosts. Transcription termination signal appropriately disposed at the 3' end of the transcribed region may also be included in the polynucleotide construct.

For secretion of the translated protein into the lumen of the endoplasmic reticulum, into the periplasmic space or into the extracellular environment, appropriate secretion signals may be incorporated into the expressed polypeptide. These signals may be endogenous to the polypeptide or they may be heterologous signals.

The polypeptide may be expressed in a modified form, such as a fusion protein, and may include not only secretion signals but also additional heterologous functional regions. Thus, for instance, a region of additional amino acids, particularly charged amino acids, may be added to the N- or C-terminus of the polypeptide to improve stability and persistence in the host cell, during purification or during subsequent handling and storage. Also, region also may be added to the polypeptide to facilitate purification. Such regions may be removed prior to final preparation of the polypeptide. The addition of peptide moieties to polypeptides to engender secretion or excretion, to improve stability or to facilitate purification, among others, are familiar and routine techniques in the art. A preferred fusion protein comprises a heterologous region from immunolglobulin that is useful to solubilize or purify polypeptides. For example, EP-A-O 464 533 (Canadian counterpart 2045869) discloses fusion proteins comprising various portions of constant region of immunoglobin molecules together with another protein or part thereof. In drug discovery, for example, proteins have been fused with antibody Fc portions for the purpose of high-throughput screening assays to identify antagonists. *See*, D. Bennett, *et al.*, *Journal of Molecular Recognition*, 8: 52-58 (1995) and K. Johanson, *et al.*, *The Journal of Biological Chemistry*, 270 (16): 9459-9471 (1995).

hybridizing amplified DNA to radiolabeled 11cb splice variant RNA or, radiolabeled 11cb splice variant antisense DNA sequences. Perfectly matched sequences can be distinguished from mismatched duplexes by RNase A digestion or by differences in melting temperatures.

Sequence differences between a reference gene and genes having mutations may also be revealed by direct DNA sequencing. In addition, cloned DNA segments may be employed as probes to detect specific DNA segments. The sensitivity of such methods can be greatly enhanced by appropriate use of PCR or other amplification methods. For example, a sequencing primer is used with double-stranded PCR product or a single-stranded template molecule generated by a modified PCR. The sequence determination is performed by conventional procedures with radiolabeled nucleotide or by automatic sequencing procedures with fluorescent-tags.

Genetic testing based on DNA sequence differences may be achieved by detection of alterations in electrophoretic mobility of DNA fragments in gels, with or without denaturing agents. Small sequence deletions and insertions can be visualized by high resolution gel electrophoresis. DNA fragments of different sequences may be distinguished on denaturing formamide gradient gels in which the mobilities of different DNA fragments are retarded in the gel at different positions according to their specific melting or partial melting temperatures. *See, e.g.*, Myers, *et al.*, *Science*, 1985, 230: 1242).

Sequence changes at specific locations may also be revealed by nuclease protection assays, such as RNase and S1 protection or the chemical cleavage method (*e.g.*, Cotton, *et al.*, *Proc. Natl. Acad. Sci., USA*, 1985, 85: 4397-4401).

Thus, the detection of a specific DNA sequence may be achieved by methods such as hybridization, RNase protection, chemical cleavage, direct DNA sequencing or the use of restriction enzymes, (*e.g.*, restriction fragment length polymorphisms ("RFLP") and Southern blotting of genomic DNA. In addition to more conventional gel-electrophoresis and DNA sequencing, mutations can also be detected by in situ analysis.

In accordance with a further aspect of the invention, there is provided a process for diagnosing or determining a susceptibility to infections such as bacterial, fungal, protozoan and viral infections, particularly infection caused by HIV-1 or HIV-2; pain; cancers; diabetes; obesity; feeding and drinking abnormalities, such as anorexia and bulimia; asthma; Parkinson's disease; both acute and congestive heart failure; hypotension; hypertension; urinary retention; osteoporosis; angina pectoris; myocardial infarction; ulcers; allergies; benign prostatic

It is then necessary to determine the differences in the cDNA or genomic sequence between affected and unaffected individuals. If a mutation is observed in some or all of the affected individuals but not in any normal individuals, then the mutation is likely to be the causative agent of the disease.

Polypeptide assays

The present invention also relates to a diagnostic assays such as quantitative and diagnostic assays for detecting levels of the 11cb splice variant protein in cells and tissues, including determination of normal and abnormal levels. Thus, for instance, a diagnostic assay in accordance with the invention for detecting over-expression of the 11cb splice variant protein compared to normal control tissue samples may be used to detect the presence of a disease/disorder such as infections, including bacterial, fungal, protozoan and viral infections, particularly infection caused by HIV-1 or HIV-2; pain; cancers; diabetes; ; feeding and drinking abnormalities, such as anorexia and bulimia; asthma; Parkinson's disease; both acute and congestive heart failure; hypotension; hypertension; urinary retention; osteoporosis; angina pectoris; myocardial infarction; ulcers; allergies; benign prostatic hypertrophy and psychotic and neurological disorders, including anxiety, schizophrenia, manic depression, delirium, dementia or severe mental retardation; and dyskinesias, such as Huntington's disease or Gilles dela Tourett's syndrome, among others. Assay techniques that can be used to determine levels of a protein, such as an 11cb splice variant protein of the present invention, in a sample derived from a host are well-known to those of skill in the art. Such assay methods include radioimmunoassays, competitive-binding assays, Western Blot analysis and ELISA assays. Among these ELISAs are frequently preferred. An ELISA assay initially comprises preparing an antibody specific to the 11cb splice variant, preferably a monoclonal antibody. In addition, a reporter antibody generally is prepared which binds to the monoclonal antibody. The reporter antibody is attached a detectable reagent such as a radioactive, fluorescent or enzymatic reagent, in this example horseradish peroxidase enzyme.

To carry out an ELISA a sample is removed from a host and incubated on a solid support, *e.g.*, a polystyrene dish, that binds the proteins in the sample. Any free protein binding sites on the dish are then covered by incubating with a non-specific protein such as bovine serum albumin. Next, the monoclonal antibody is incubated in the dish during which time the monoclonal antibodies attach to any 11cb splice variant proteins attached to the polystyrene dish. Unbound monoclonal antibody is washed out with buffer. The reporter antibody linked to horseradish peroxidase is placed in the dish resulting in binding of the reporter antibody to any

may be used to express humanized antibodies to immunogenic polypeptide products of this invention.

The above-described antibodies may be employed to isolate or to identify clones expressing the polypeptide or purify the polypeptide of the present invention by attachment of the antibody to a solid support for isolation and/or purification by affinity chromatography.

Antibodies against the 11cb splice variant may also be employed to inhibit infections, such as bacterial, fungal, protozoan and viral infections, particularly infection caused by HIV-1 or HIV-2; pain; cancers; diabetes; obesity; feeding and drinking abnormalities, such as anorexia and bulimia; asthma; Parkinson's disease; both acute and congestive heart failure; hypotension; hypertension; urinary retention; osteoporosis; angina pectoris; myocardial infarction; ulcers; allergies; benign prostatic hypertrophy and psychotic and neurological disorders, including anxiety, schizophrenia, manic depression, delirium, dementia or severe mental retardation, and dyskinesias, such as Huntington's disease or Gilles dela Tourett's syndrome, among others.

11cb Splice Variant Binding Molecules and Assays

The 11cb splice variant can be used to isolate proteins which interact with it; this interaction can be a target for interference. Inhibitors of protein-protein interactions between the 11cb splice variant and other factors could lead to the development of pharmaceutical agents for the modulation of 11cb splice variant activity.

Thus, this invention also provides a method for identification of binding molecules to the 11cb splice variant. Genes encoding proteins for binding molecules to the 11cb splice variant can be identified by numerous methods known to those of skill in the art, for example, ligand panning and FACS sorting. Such methods are described in many laboratory manuals such as, for instance, Coligan, *et al.*, *Current Protocols in Immunology* 1: Chapter 5 (1991) and Rivett, A. J., *Biochem.* (1993), 291: 1-10.

For example, the yeast two-hybrid system provides methods for detecting the interaction between a first test protein and a second test protein, *in vivo*, using reconstitution of the activity of a transcriptional activator. The method is disclosed in U.S. Patent No. 5,283,173; reagents are available from Clontech and Stratagene. Briefly, 11cb splice variant cDNA is fused to a Gal4 transcription factor DNA binding domain and expressed in yeast cells. cDNA library members obtained from cells of interest are fused to a transactivation domain of Gal4. cDNA clones which express proteins which can interact with the 11cb splice variant will lead to reconstitution of Gal4 activity and transactivation of expression of a reporter gene such as Gal1-lacZ.

SDS-PAGE. Binding partner primary amino acid sequence data are obtained by microsequencing. Optionally, the cells can be treated with agents that induce a functional response such as tyrosine phosphorylation of cellular proteins. An example of such an agent would be a growth factor or cytokine such as interleukin-2.

Another alternative method is immunoaffinity purification. Recombinant 11cb splice variant is incubated with labeled or unlabeled cell extracts and immunoprecipitated with anti-11cb splice variant antibodies. The immunoprecipitate is recovered with protein A-Sepharose and analyzed by SDS-PAGE. Unlabelled proteins are labeled by biotinylation and detected on SDS gels with streptavidin. Binding partner proteins are analyzed by microsequencing. Further, standard biochemical purification steps known to those skilled in the art may be used prior to microsequencing.

Yet another alternative method is screening of peptide libraries for binding partners. Recombinant tagged or labeled 11cb splice variant is used to select peptides from a peptide or phosphopeptide library which interact with the 11cb splice variant. Sequencing of the peptides leads to identification of consensus peptide sequences which might be found in interacting proteins.

The 11cb splice variant binding partners identified by any of these methods or other methods which would be known to those of ordinary skill in the art, as well as those putative binding partners discussed above, can be used in the assay method of the invention. Assaying for the presence of the 11cb splice variant/binding partner complex are accomplished by, for example, the yeast two-hybrid system, ELISA or immunoassays using antibodies specific for the complex. In the presence of test substances which interrupt or inhibit formation of the 11cb splice variant/binding partner interaction, a decreased amount of complex will be determined relative to a control lacking the test substance.

Assays for free 11cb splice variant or binding partner are accomplished by, for example, ELISA or immunoassay using specific antibodies or by incubation of radiolabeled 11cb splice variant with cells or cell membranes followed by centrifugation or filter separation steps. In the presence of test substances which interrupt or inhibit formation of the 11cb splice variant/binding partner interaction, an increased amount of free 11cb splice variant or free binding partner is determined relative to a control lacking the test substance.

Polypeptides of the invention also can be used to assess 11cb splice variant binding capacity of 11cb splice variant binding molecules in cells or in cell-free preparations.

Another method involves screening for compounds which are antagonists, and thus inhibit activation of the receptor polypeptide of the present invention by determining inhibition of binding of labeled ligand, such as MCH, to cells which have the receptor on the surface thereof, or cell membranes containing the receptor. Such a method involves transfecting a eukaryotic cell with DNA encoding the 11cb splice variant such that the cell expresses the receptor on its surface. The cell is then contacted with a potential antagonistin the presence of a labeled form of a ligand, such as MCH. The ligand can be labeled, *e.g.*, by radioactivity. The amount of labeled ligand bound to the receptors is measured, *e.g.*, by measuring radioactivity associated with transfected cells or membrane from these cells. If the compound binds to the receptor, the binding of labeled ligand to the receptor is inhibited as determined by a reduction of labeled ligand which binds to the receptors. This method is called binding assay.

Another such screening procedure involves the use of mammalian cells which are transfected to express the receptor of interest. The cells are loaded with an indicatator dye that produces a fluorescent signal when bound to calcium, and the cells are contacted with a test substance and a receptor agonist, such as MCH. Any change in fluorescent signal is measured over a defined period of time using, for example, a fluorescence spectrophotometer or a fluorescence imaging plate reader. A change in the fluorescence signal pattern generated by the ligand indicates that a compound is a potential antagonist (or agonist) for the receptor.

Another such screening procedure involves use of mammalian cells which are transfected to express the receptor of interest, and which are also transfected with a reporter gene construct that is coupled to activation of the receptor (for example, luciferase or beta-galactosidase behind an appropriate promoter). The cells are contacted with a test substance and the receptor agonist, such as MCH, and the signal produced by the reporter gene is measured after a defined period of time. The signal can be measured using a luminometer, spectrophotometer, fluorimeter, or other such instrument appropriate for the specific reporter construct used. Inhibition of the signal generated by the ligand indicates that a compound is a potential antagonist for the receptor.

Another such screening technique for antagonists or agonits involves introducing RNA encoding the 11cb splice variant into *Xenopus* oocytes to transiently or stably express the receptor. The receptor oocytes are then contacted with the receptor ligand, such as MCH, and a compound to be screened. Inhibition or activation of the receptor is then determined by detection of a signal, such as, cAMP, calcium, proton, or other ions.

4700-4709, 1996). This provides a rapid direct growth selection (e.g, using the *FUS1-HIS3* reporter) for surrogate peptide agonists that activate characterized or orphan receptors. Alternatively, yeast cells that functionally express human (or mammalian) G-protein coupled receptors linked to a reporter gene readout (e.g., *FUS1-LacZ*) can be used as a platform for high-throughput screening of known ligands, fractions of biological extracts and libraries of chemical compounds for either natural or surrogate ligands. Functional agonists of sufficient potency (whether natural or surrogate) can be used as screening tools in yeast cell-based assays for identifying G-protein coupled receptor antagonists. For this purpose, the yeast system offers advantages over mammalian expression systems due to its ease of utility and null receptor background (lack of endogenous G-protein coupled receptors) which often interferes with the ability to identify agonists or antagonists.

The present invention also provides a method for determining whether a ligand not known to be capable of binding to an 11cb splice variant receptor can bind to such receptor which comprises contacting a mammalian cell which expresses an 11cb splice variant receptor with the ligand such as MCH under conditions permitting binding of candidate ligands to the 11cb splice variant receptor, and detecting the presence of a candidate ligand which binds to the receptor thereby determining whether the ligand binds to the 11cb splice variant receptor. The systems hereinabove described for determining agonists and/or antagonists may also be employed for determining ligands which bind to the receptor.

Examples of potential 11cb splice variant receptor antagonists include antibodies or, in some cases, oligonucleotides, which bind to the receptor but do not elicit a second messenger response such that the activity of the receptor is prevented.

Potential antagonists also include proteins which are closely related to the ligand of the 11cb splice variant receptor, *i.e.* a fragment of the ligand, which have lost biological function and when binding to the 11cb splice variant receptor, elicit no response.

Thus in another aspect, the present invention relates to a screening kit for identifying agonists, antagonists, and ligands for 11cb splice variant polypeptides, which comprises:

(a) a 11cb splice variant polypeptide, preferably that of SEQ ID NO:2; and further preferably comprises labeled or unlabeled MCH;

(b) a recombinant cell expressing a 11cb splice variant polypeptide, preferably that of SEQ ID NO:2; and further preferably comprises labeled or unlabeled MCH; or

protozoan and viral infections, particularly infection caused by HIV-1 or HIV-2; pain; cancers; diabetes; obesity; feeding and drinking abnormalities, such as anorexia and bulimia; asthma; Parkinson's disease; both acute and congestive heart failure; hypotension; hypertension; urinary retention; osteoporosis; angina pectoris; myocardial infarction; ulcers; allergies; benign prostatic hypertrophy and psychotic and neurological disorders, including anxiety, schizophrenia, manic depression, delirium, dementia or severe mental retardation; or dyskinesias, such as Huntington's disease or Gilles dela Tourett's syndrome, among others.

Antagonists for the 11cb splice variant may be employed for a variety of therapeutic and prophylactic purposes for such diseases or disorders as infections, including bacterial, fungal, protozoan and viral infections, particularly infection caused by HIV-1 or HIV-2; pain; cancers; diabetes; obesity; feeding and drinking abnormalities, such as anorexia and bulimia; asthma; Parkinson's disease; both acute and congestive heart failure; hypotension; hypertension; urinary retention; osteoporosis; angina pectoris; myocardial infarction; ulcers; allergies; benign prostatic hypertrophy and psychotic and neurological disorders, including anxiety, schizophrenia, manic depression, delirium, dementia or severe mental retardation; or dyskinesias, such as Huntington's disease or Gilles dela Tourett's syndrome, among others.

This invention additionally provides a method of treating an abnormal condition related to an excess of 11cb splice variant activity which comprises administering to a subject the inhibitor compounds (antagonists) as hereinabove described along with a pharmaceutically acceptable carrier in an amount effective to inhibit activation by blocking binding of ligands to the 11cb splice variant, or by inhibiting a second signal, and thereby alleviating the abnormal conditions.

The invention also provides a method of treating abnormal conditions related to an under-expression of 11cb splice variant activity which comprises administering to a subject a therapeutically effective amount of a compound which activates the receptor polypeptide of the present invention (agonists) as described above in combination with a pharmaceutically acceptable carrier, to thereby alleviate the abnormal conditions.

Compositions and Kits

The soluble form of the 11cb splice variant, and compounds which activate or inhibit such receptor, may be employed in combination with a suitable pharmaceutical carrier. Such compositions comprise a therapeutically effective amount of the polypeptide or compound, and a pharmaceutically acceptable carrier or excipient. Such a carrier includes but is not limited to

use of a retroviral plasmid vector containing RNA encoding a polypeptide of the present invention. Such methods are well-known in the art and their use in the present invention will be apparent from the teachings herein.

Similarly, cells may be engineered in vivo for expression of a polypeptide in vivo by procedures known in the art. For example, a polynucleotide of the invention may be engineered for expression in a replication defective retroviral vector, as discussed above. The retroviral expression construct may then be isolated and introduced into a packaging cell is transduced with a retroviral plasmid vector containing RNA encoding a polypeptide of the present invention such that the packaging cell now produces infectious viral particles containing the gene of interest. These producer cells may be administered to a patient for engineering cells in vivo and expression of the polypeptide in vivo. These and other methods for administering a polypeptide of the present invention should be apparent to those skilled in the art from the teachings of the present invention.

Retroviruses from which the retroviral plasmid vectors herein above mentioned may be derived include, but are not limited to, Moloney Murine Leukemia Virus, Spleen Necrosis Virus, Rous Sarcoma Virus, Harvey Sarcoma Virus, Avian Leukosis Virus, Gibbon Ape Leukemia Virus, Human Immunodeficiency Virus, Adenovirus, Myeloproliferative Sarcoma Virus, and Mammary Tumor Virus. In a preferred embodiment, the retroviral plasmid vector is derived from Moloney Murine Leukemia Virus.

Such vectors well include one or more promoters for expressing the polypeptide. Suitable promoters which may be employed include, but are not limited to, the retroviral LTR; the SV40 promoter; and the human cytomegalovirus (CMV) promoter described in Miller, *et al.*, *Biotechniques* 7: 980-990 (1989), or any other promoter (*e.g.*, cellular promoters such as eukaryotic cellular promoters including, but not limited to, the histone, RNA polymerase III, and ß-actin promoters). Other viral promoters which may be employed include, but are not limited to, adenovirus promoters, thymidine kinase (TK) promoters, and B19 parvovirus promoters. The selection of a suitable promoter will be apparent to those skilled in the art from the teachings contained herein.

The nucleic acid sequence encoding the polypeptide of the present invention will be placed under the control of a suitable promoter. Suitable promoters which may be employed include, but are not limited to, adenoviral promoters, such as the adenoviral major late promoter; or heterologous promoters, such as the cytomegalovirus (CMV) promoter; the respiratory

Sambrook and numerous other references such as, for instance, by Goeddel, *et al.*, *Nucleic Acids Res.* 8: 4057.

Unless described otherwise, ligations are accomplished using standard buffers, incubation temperatures and times, approximately equimolar amounts of the DNA fragments to be ligated and approximately 10 units of T4 DNA ligase ("ligase") per 0.5 μg of DNA.

Example 1 - Cloning of the 5'-end of the Human 11cb Splice Variant

The 5' end of the human 11cb splice variant was amplified by using the following primers and conditions on DNA from Human Whole Brain, purified from Life Technology's plasmid libraries.

The outside primers used were:

5' Vector-specific primer: 5' GCT ATT TAG GTG ACA CTA TAG AAG GTA CG 3'

(SEQ ID NO: 3); and

3' Gene-specific primer: 5' CGA GAG GTT GAT GAT GAA GAT GTC 3'

(SEQ ID NO: 4).

The following was used in a 50 microliter reaction volume: 10X Taq polymerase buffer, 200μM dNTP, 5% glycerol, 50 picomoles of each primer, 100 nanograms of plasmid DNA from Human Whole Brain, purified from Life Technology's plasmid libraries, 1:10 by volume mixture of Taq polymerase and Pfu polymerase.

The following PCR program was then used:

1 cycle at 94°C for 5 minutes;

25 cycles at 94°C for 1 minute;

25 cycles at 55°C for 1 minute;

25 cycles at 72°C for 2 minutes; and

1 cycle at 72°C for 8.5 minutes.

The nested or inside primers used were:

5' Vector-specific primer: 5' GGT GAC ACT ATA GAA GGT ACG 3'

(SEQ ID NO: 5) and

3' Gene-specific primer: 5' GCA GAT GGT GCC GAA CAC CGA AGG 3'

2. Microphysiometer assay

Screening of these banks is accomplished using a microphysiometer (commercially available from, *e.g.*, Molecular Devices, Ltd.). For example activation of secondary messenger systems results in the extrusion of small amounts of acid from a cell, formed largely as a result of increased metabolic activity required to fuel the intracellular signaling process. The pH changes in the media surrounding the cell are small and detectable by the microphysiometer. Thus activation of any receptor which is coupled to an energy utilizing intracellular signaling pathway (*e.g.*, andy G-protein coupled receptor) may be detectable.

3. Calcium Assay

Receptors stably expressed in HEK 293 cells can demonstrate a robust calcium response to agonists with the appropriate rank order and potency. Basal calcium levels in the HEK 293 cells in receptor-transfected or vector control cells is in the normal 100 nM to 200 nM range. HEK 293 cells expressing recombinant receptors are loaded with fura 2 and in a single day > 150 selected ligands are evaluated for agonist-induced calcium mobilization. Agonists presenting a transient calcium mobilization are tested in vector control cells to determine if the calcium response was unique to the transfected receptor cells. When a unique agonist-induced response is identified, the response is reproduced in a separate group of cells and then pharmacologically characterized with concentration response curves for the effective and related ligands.

Example 2 - Northern Blot Analysis

The northern blots used were purchased from Clontech. The transcript size was approximately 2.4kb, and a transcript band was observed in whole brain, amygdala, caudate nucleus, corpus callosum, hippocampus, substantia nigra, subthalamic nucleus, thalamus, heart, and liver. Conversely, no transcript bands were detected by northern blot analysis in the following tissues: placenta, lung, skeletal muscle, kidney, or pancreas.

Example 3

HEK 293 cells transiently transfected with 11cb splice variant responded with a robust calcium mobilization response to the 19 amino acid peptide melanin-concentrating hormone (MCH) with amino acid sequence of H-Asp-Phe-Asp-Met-Leu-Arg-Cys-Met-Leu-Gly-Arg-Val-Tyr-Arg-Pro-Cys-Trp-Gln-Val-OH (SEQ ID NO: 7). Thus, it has now been found that MCH is a ligand for 11cb splice variant. MCH is a peptide present in the brain of veterbrates and functions as a pigment cell agonist in fish, regulating melanocyte proliferation and melanin synthesis. In

What is claimed is:

1. An isolated polynucleotide comprising a member selected from the group consisting of:

(a) a polynucleotide that is at least 91% identical to a polynucleotide encoding a polypeptide comprising amino acids of SEQ ID NO: 2;

(b) a polynucleotide which by virtue of the redundancy of the genetic code, encodes the same amino acids of SEQ ID NO: 2;

(c) a polynucleotide which is complementary to the polynucleotide of (a) or (b); and

(d) a polynucleotide comprising at least 15 contiguous bases of the polynucleotide of (a), (b) or (c).

2. The polynucleotide of Claim 1 wherein the polynucleotide is DNA.

3. The polynucleotide of Claim 1 wherein the polynucleotide is RNA.

4. The polynucleotide of Claim 2 comprising nucleotides set forth in SEQ ID NO: 1.

5. The polynucleotide of Claim 2 which encodes a polypeptide comprising amino acids of SEQ ID NO: 2.

6. A vector comprising the DNA of Claim 2.

7. A host cell comprising the vector of Claim 6.

8. A process for producing an 11cb splice variant polypeptide comprising: culturing a host of claim 7 in a medium and under conditions sufficient for the expression of said polypeptide and recovering the expressed polypeptide.

18. A process for diagnosing a disease or a susceptibility to a disease related to expression of the polypeptide of claim 10 comprising determining a mutation in the nucleic acid sequence encoding said polypeptide.

19. A diagnostic process comprising analyzing for the presence of the polypeptide of claim 11 in a sample derived from a host.

20. A method for identifying agonist or antagonist of a polypeptide of claim 10 which comprises:

contacting a cell expressing on the surface thereof the polypeptide, said polypeptide being associated with a second component capable of providing a detectable signal in response to the binding of a compound to said polypeptide, with a compound to be screened under conditions to permit binding to the polypeptide; and

determining whether the compound binds to and activates or inhibits the polypeptide by measuring the level of a signal generated from the interaction of the compound with the polypeptide.

21. A method of claim 20 which further comprises conducting the identification of agonist or antagonist in the presence of labeled or unlabeled MCH.

22. A method for identifying agonist or antagonist of a polypeptide of claim 10 which comprises:

determining the inhibition of binding of a ligand to cells which have the polypeptide on the surface thereof, or to cell membranes containing the polypeptide, in the presence of a candidate compound under conditions to permit binding to the polypeptide, and determining the amount of ligand bound to the polypeptide, such that a compound capable of causing reduction of binding of a ligand is an agonist or antagonist.

23. A method of claim 22 in which a ligand is labeled or unlabeled MCH.

Nucleotide sequence of the human 11cb splice variant (SEQ ID NO: 1).

```
   1  GGTGACACTA TAGAAGGTAC GCCTGCAGGT ACCGGTCCGG AATTCCCGGG
  51  TCGACCCACG CGTCCGGGAG GGCAGTTGGG CTTGGAGGCG GCAGCGGCTG
 101  CCAGGCTACG GAGGAAGACC CCCTTCCCGA CTGCGGGGCT TGCGCTCCGG
 151  GACAAGGTGG CAGGCGCTGG AGGCTGCCGC AGCCTGCGTG GGTGGAGGGG
 201  AGCTCAGCTC GGTTGTGGGA GCAGGCGACC GGCACTGGCT GGATGGACCT
 251  GGAAGCCTCG CTGCTGCCCA CTGGTCCCAA TGCCAGCAAC ACCTCTGATG
 301  GCCCCGATAA CCTCACTTCG GCAGGATCAC CTCCTCGCAC GGGGAGCATC
 351  TCCTACATCA ACATCATCAT GCCTTCGGTG TTCGGCACCA TCTGCCTCCT
 401  GGGCATCATC GGGAACTCCA CGGTCATCTT CGCGGTCGTG AAGAAGTCCA
 451  AGCTGCACTG GTGCAACAAC GTCCCCGACA TCTTCATCAT CAACCTCTCG
 501  GTAGTAGATC TCCTCTTTCT CCTGGGCATG CCCTTCATGA TCCACCAGCT
 551  CATGGGCAAT GGGGTGTGGC ACTTTGGGGA GACCATGTGC ACCCTCATCA
 601  CGGCCATGGA TGCCAATAGT CAGTTCACCA GCACCTACAT CCTGACCGCC
 651  ATGGCCATTG ACCGCTACCT GGCCACTGTC CACCCCATCT CTTCCACGAA
 701  GTTCCGGAAG CCCTCTGTGG CCACCCTGGT GATCTGCCTC CTGTGGGCCC
 751  TCTCCTTCAT CAGCATCACC CCTGTGTGGC TGTATGCCAG ACTCATCCCC
 801  TTCCCAGGAG GTGCAGTGGG CTGCGGCATA CGCCTGCCCA ACCCAGACAC
 851  TGACCTCTAC TGGTTCACCC TGTACCAGTT TTTCCTGGCC TTTGCCCTGC
 901  CTTTTGTGGT CATCACAGCC GCATACGTGA GGATCCTGCA GCGCATGACG
 951  TCCTCAGTGG CCCCCGCCTC CCAGCGCAGC ATCCGGCTGC GGACAAAGAG
```

FIG. 1

```
1001  GGTGACCCGC ACAGCCATCG CCATCTGTCT GGTCTTCTTT GTGTGCTGGG
1051  CACCCTACTA TGTGCTACAG CTGACCCAGT TGTCCATCAG CCGCCCGACC
1101  CTCACCTTTG TCTACTTATA CAATGCGGCC ATCAGCTTGG GCTATGCCAA
1151  CAGCTGCCTC AACCCCTTTG TGTACATCGT GCTCTGTGAG ACGTTCCGCA
1201  AACGCTTGGT CCTGTCGGTG AAGCCTGCAG CCCAGGGGCA GCTTCGCGCT
1251  GTCAGCAACG CTCAGACGGC TGACGAGGAG AGGACAGAAA GCAAAGGCAC
1301  CTGATACTTC CCCTGCCACC CTGCACACCT CCAAGTCAGG GCACCACAAC
1351  ACGCCACCGG GAGAGATGCT CTCGTGCCGA ATTCC
```

FIG. 1A

Amino Acid sequence of the Human 11cb Splice Variant (SEQ ID NO: 2).

```
  1  MDLEASLLPT GPNASNTSDG PDNLTSAGSP PRTGSISYIN IIMPSVFGTI
 51  CLLGIIGNST VIFAVVKKSK LHWCNNVPDI FIINLSVVDL LFLLGMPFMI
101  HQLMGNGVWH FGETMCTLIT AMDANSQFTS TYILTAMAID RYLATVHPIS
151  STKFRKPSVA TLVICLLWAL SFISITPVWL YARLIPFPGG AVGCGIRLPN
201  PDTDLYWFTL YQFFLAFALP FVVITAAYVR ILQRMTSSVA PASQRSIRLR
251  TKRVTRTAIA ICLVFFVCWA PYYVLQLTQL SISRPTLTFV YLYNAAISLG
301  YANSCLNPFV YIVLCETFRK RLVLSVKPAA QGQLRAVSNA QTADEERTES
351  KGT
```

FIG. 1B

SEQUENCE LISTING

(1) GENERAL INFORMATION

(i) APPLICANT: SMITHKLINE BEECHAM CORPORATION
SMITHKLINE BEECHAM p.l.c.

(ii) TITLE OF THE INVENTION: A METHOD OF FINDING AGONIST AND ANTAGONIST TO HUMAN 11CB SPLICE VARIANT

(iii) NUMBER OF SEQUENCES: 7

(iv) CORRESPONDENCE ADDRESS:
(A) ADDRESSEE: Ratner & Prestia
(B) STREET: P.O. Box 980
(C) CITY: Valley Forge
(D) STATE: PA
(E) COUNTRY: USA
(F) ZIP: 19482

(v) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Diskette
(B) COMPUTER: IBM Compatible
(C) OPERATING SYSTEM: DOS
(D) SOFTWARE: FastSEQ for Windows Version 2.0

(vi) CURRENT APPLICATION DATA:
(A) APPLICATION NUMBER:
(B) FILING DATE:
(C) CLASSIFICATION: UNKNOWN

(vii) PRIOR APPLICATION DATA:
(A) APPLICATION NUMBER: 08/984,288
(B) FILING DATE: 11-DEC-1997

(A) APPLICATION NUMBER: 60/073,747
(B) FILING DATE: 05-FEB-1998

(A) APPLICATION NUMBER: 09/060,504
(B) FILING DATE: 15-APR-1998

(viii) ATTORNEY/AGENT INFORMATION:
(A) NAME: Prestia, Paul F
(B) REGISTRATION NUMBER: 23,031
(C) REFERENCE/DOCKET NUMBER: GP-50003-1

(ix) TELECOMMUNICATION INFORMATION:
(A) TELEPHONE: 610-407-0700
(B) TELEFAX: 610-407-0700
(C) TELEX: 846169

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1385 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

```
GGTGACACTA TAGAAGGTAC GCCTGCAGGT ACCGGTCCGG AATTCCCGGG TCGACCCACG     60
CGTCCGGGAG GGCAGTTGGG CTTGGAGGCG GCAGCGGCTG CCAGGCTACG GAGGAAGACC    120
CCCTTCCCGA CTGCGGGGCT TGCGCTCCGG GACAAGGTGG CAGGCGCTGG AGGCTGCCGC    180
AGCCTGCGTG GGTGGAGGGG AGCTCAGCTC GGTTGTGGGA GCAGGCGACC GGCACTGGCT    240
GGATGGACCT GGAAGCCTCG CTGCTGCCCA CTGGTCCCAA TGCCAGCAAC ACCTCTGATG    300
GCCCCGATAA CCTCACTTCG GCAGGATCAC CTCCTCGCAC GGGGAGCATC TCCTACATCA    360
ACATCATCAT GCCTTCGGTG TTCGGCACCA TCTGCCTCCT GGGCATCATC GGGAACTCCA    420
CGGTCATCTT CGCGGTCGTG AAGAAGTCCA AGCTGCACTG GTGCAACAAC GTCCCCGACA    480
TCTTCATCAT CAACCTCTCG GTAGTAGATC TCCTCTTTCT CCTGGGCATG CCCTTCATGA    540
TCCACCAGCT CATGGGCAAT GGGGTGTGGC ACTTTGGGGA GACCATGTGC ACCCTCATCA    600
CGGCCATGGA TGCCAATAGT CAGTTCACCA GCACCTACAT CCTGACCGCC ATGGCCATTG    660
ACCGCTACCT GGCCACTGTC CACCCCATCT CTTCCACGAA GTTCCGGAAG CCCTCTGTGG    720
CCACCCTGGT GATCTGCCTC CTGTGGGCCC TCTCCTTCAT CAGCATCACC CCTGTGTGGC    780
TGTATGCCAG ACTCATCCCC TTCCCAGGAG GTGCAGTGGG CTGCGGCATA CGCCTGCCCA    840
ACCCAGACAC TGACCTCTAC TGGTTCACCC TGTACCAGTT TTTCCTGGCC TTTGCCCTGC    900
CTTTTGTGGT CATCACAGCC GCATACGTGA GGATCCTGCA GCGCATGACG TCCTCAGTGG    960
CCCCCGCCTC CCAGCGCAGC ATCCGGCTGC GGACAAAGAG GGTGACCCGC ACAGCCATCG   1020
CCATCTGTCT GGTCTTCTTT GTGTGCTGGG CACCCTACTA TGTGCTACAG CTGACCCAGT   1080
TGTCCATCAG CCGCCCGACC CTCACCTTTG TCTACTTATA CAATGCGGCC ATCAGCTTGG   1140
GCTATGCCAA CAGCTGCCTC AACCCCTTTG TGTACATCGT GCTCTGTGAG ACGTTCCGCA   1200
AACGCTTGGT CCTGTCGGTG AAGCCTGCAG CCCAGGGGCA GCTTCGCGCT GTCAGCAACG   1260
CTCAGACGGC TGACGAGGAG AGGACAGAAA GCAAAGGCAC CTGATACTTC CCCTGCCACC   1320
CTGCACACCT CCAAGTCAGG GCACCACAAC ACGCCACCGG GAGAGATGCT CTCGTGCCGA   1380
ATTCC                                                               1385
```

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 353 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

```
Met Asp Leu Glu Ala Ser Leu Leu Pro Thr Gly Pro Asn Ala Ser Asn
 1               5                  10                  15
Thr Ser Asp Gly Pro Asp Asn Leu Thr Ser Ala Gly Ser Pro Pro Arg
            20                  25                  30
Thr Gly Ser Ile Ser Tyr Ile Asn Ile Ile Met Pro Ser Val Phe Gly
        35                  40                  45
Thr Ile Cys Leu Leu Gly Ile Ile Gly Asn Ser Thr Val Ile Phe Ala
    50                  55                  60
Val Val Lys Lys Ser Lys Leu His Trp Cys Asn Asn Val Pro Asp Ile
65                  70                  75                  80
Phe Ile Ile Asn Leu Ser Val Val Asp Leu Leu Phe Leu Leu Gly Met
                85                  90                  95
Pro Phe Met Ile His Gln Leu Met Gly Asn Gly Val Trp His Phe Gly
            100                 105                 110
Glu Thr Met Cys Thr Leu Ile Thr Ala Met Asp Ala Asn Ser Gln Phe
        115                 120                 125
Thr Ser Thr Tyr Ile Leu Thr Ala Met Ala Ile Asp Arg Tyr Leu Ala
```

```
    130                 135                 140

Thr Val His Pro Ile Ser Ser Thr Lys Phe Arg Lys Pro Ser Val Ala
145                 150                 155                 160
Thr Leu Val Ile Cys Leu Leu Trp Ala Leu Ser Phe Ile Ser Ile Thr
                165                 170                 175
Pro Val Trp Leu Tyr Ala Arg Leu Ile Pro Phe Pro Gly Gly Ala Val
            180                 185                 190
Gly Cys Gly Ile Arg Leu Pro Asn Pro Asp Thr Asp Leu Tyr Trp Phe
        195                 200                 205
Thr Leu Tyr Gln Phe Phe Leu Ala Phe Ala Leu Pro Phe Val Val Ile
    210                 215                 220
Thr Ala Ala Tyr Val Arg Ile Leu Gln Arg Met Thr Ser Ser Val Ala
225                 230                 235                 240
Pro Ala Ser Gln Arg Ser Ile Arg Leu Arg Thr Lys Arg Val Thr Arg
                245                 250                 255
Thr Ala Ile Ala Ile Cys Leu Val Phe Phe Val Cys Trp Ala Pro Tyr
            260                 265                 270
Tyr Val Leu Gln Leu Thr Gln Leu Ser Ile Ser Arg Pro Thr Leu Thr
        275                 280                 285
Phe Val Tyr Leu Tyr Asn Ala Ala Ile Ser Leu Gly Tyr Ala Asn Ser
    290                 295                 300
Cys Leu Asn Pro Phe Val Tyr Ile Val Leu Cys Glu Thr Phe Arg Lys
305                 310                 315                 320
Arg Leu Val Leu Ser Val Lys Pro Ala Ala Gln Gly Gln Leu Arg Ala
                325                 330                 335
Val Ser Asn Ala Gln Thr Ala Asp Glu Glu Arg Thr Glu Ser Lys Gly
            340                 345                 350
Thr
```

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 29 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

GCTATTTAGG TGACACTATA GAAGGTACG 29

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 24 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

CGAGAGGTTG ATGATGAAGA TGTC 24

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 21 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

GGTGACACTA TAGAAGGTAC G 21

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 24 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

GCAGATGGTG CCGAACACCG AAGG 24

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 19 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

Asp Phe Asp Met Leu Arg Cys Met Leu Gly Arg Val Tyr Arg Pro Cys
1 5 10 15
Trp Gln Val